Naval Medical Research Institute
Bethesda, MD 20814-5055

NMRI-33 November 1988

STRATEGY FOR ISOLATION OF GENE ACTIVATION FACTORS

HENRY CHANG

Naval Medical Research
and Development Command
Bethesda, Maryland 20814-5044

Department of the Navy
Naval Medical Command
Washington, D.C. 20372-5210

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b RESTRICTIVE MARKINGS:
2a SECURITY CLASSIFICATION AUTHORITY:	3 DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution is unlimited
2b DECLASSIFICATION/DOWNGRADING SCHEDULE:	
4 PERFORMING ORGANIZATION REPORT NUMBER(S): NMRI 88-33	5 MONITORING ORGANIZATION REPORT NUMBER(S):
6a NAME OF PERFORMING ORGANIZATION: Naval Medical Research Institute; 6b OFFICE SYMBOL (If applicable):	7a NAME OF MONITORING ORGANIZATION: Naval Medical Command
6c ADDRESS (City, State, and ZIP Code): Bethesda, MD 20814-5055	7b ADDRESS (City, State, and ZIP Code): Department of the Navy, Washington, DC 20372-5120
8a NAME OF FUNDING SPONSORING ORGANIZATION: Naval Medical Research & Development Command; 8b OFFICE SYMBOL (If applicable):	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER:
8c ADDRESS (City, State, and ZIP Code): Bethesda, MD 20814-5000	10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. 62233N; PROJECT NO. MM33C30; TASK NO. 007; WORK UNIT ACCESSION NO. DN377031

11 TITLE (Include Security Classification): Strategy for isolation of gene activation factors (U)

12. PERSONAL AUTHOR(S): Henry Chang

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical Report	FROM 7/87 TO 8/88	October 1988	19

16. SUPPLEMENTARY NOTATION:

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	gene regulation; trans-activating factors

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

How genes regulate one another has been a central problem in biology. One mechanism is that certain genes produce "trans-activating factors," which stimulate others, but the former have been difficult to isolate since they may be distant from the latter. This report describes a possible solution to this problem.

Recombinant DNA technology can be used to construct a vector composed of gene regulatory (e.g. promoter) sequences fused inframe to a reporter gene, plus a selectable marker. When this vector is placed into permissive host cells, transfection of additional DNAs (syngeneic with the source of the regulatory sequences) may serve to activate the reporter gene. These new DNAs, which encode trans-activating factors, then can be rescued with species-specific or linked probes. (AW)

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: Mrs. Phyllis Blum, Information Services Div.	22b. TELEPHONE (Include Area Code): (202) 295-2188; 22c. OFFICE SYMBOL: ISD/RSD/NMRI

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

Table of Contents

INTRODUCTION

The development of a living organism depends on an orderly process whereby cells, and ultimately their genes, communicate with one another. What causes different parts of the genome to express themselves during the life cycle, and only in specific tissues, has been a fundamental problem since the inception of Biology.

Even with the advent of recombinant DNA technology, this puzzle has been difficult to solve. Many different methods have been described in a recent review (1). Putative regulatory sequences (e.g. promoters, enhancers) can be cloned (2), mutated (3), and used to bind nuclear extracts for transcription assays(4) or for the detection of protein-DNA complexes by their retarded migration on gels (5). In addition, blotting (6), DNAse hypersensitivity (7), and "foot-printing" studies (8) can be done, but they do not indicate if the factors involved are functionally active. Also, DNA-binding proteins can be enriched by recognition-site affinity chromatography (9), their amino acids sequenced, and oligonucleotides probes generated to find their corresponding genes, but these steps are laborious and problematic because binding affinities may overlap to preclude purification. More directly, clones which bind DNA can be identified in bacteriophage expression libraries with specific double-stranded, non-denatured probes (10), but again, one would have to prove the function of the proteins they encode separately. Thus, it has been difficult to develop a general approach to isolate genes which regulate other genes in eukaryotic cells.

The method proposed here depends on standard molecular biology techniques arranged in a unique way to favor the identification of cells which contain a selectable reporter gene by complementation with sequences from a DNA library. For illustration, an attempt will be made to find regulators of the human fetal globin gene, which is switched on and off during ontogeny, and which may be useful in the treatment of certain hematologic disorders such as sickle cell anemia and thalassemia. The work described in this preliminary report is on-going, but the discussion of the various decision points in the strategy should be valuable to further progress in the field.

METHODS

Materials. Restriction enzymes, DNA polymerase I, Klenow fragment, T4 DNA polymerase and ligase were purchased from Bethesda Research laboratories, New England Biolabs, or Pharmacia, and used under manufacturers' conditions. Radioactive nucleotides and DNA sequencing kits were from Amersham and New England Nuclear.

Plasmid Construction. Standard molecular biology techniques were used (11) in accordance with NIH Biosafety Guidelines. The fetal globin promoter was a 1.5 Kb Bam HI-Bgl I fragment derived from a 7.2 Kb EcoR I insert which had been cloned into pUC-9 by D. Bodine and T. Ley of the NHLBI, National Institutes of Health, Bethesda, MD. The sequence was spliced to the Eag I (Xma III) site of pSV2Neo (12) (from S. Segal, NCI, NIH) as shown in Fig. 1 to create an open reading frame. Then the SV40 promoter was deleted by digestion with Acc I and Bgl II, re-ligation, and transformation of HB101. The Nde I-EcoR I fragment which contained the fused promoter-reporter gene was inserted into the EcoR I site of pSV2Gpt (13) supplied by B. Howard of NCI, NIH, as shown in Fig. 2.

DNA transfer. Electroporation was carried out with the Progenetor (Hoefer) or the Gene Pulser (Biorad) on mouse erythroleukemia cells (Friend clone 745) donated by I. Kirsch of NCI, NIH. Cells were grown in RPMI medium (Hazleton) with 10% fetal calf serum (GIBCO) and penicillin/streptomycin, supplemented with reagents for Gpt selection (Sigma) or G-418 (Geneticin, GIBCO) for Neo selection.

The complementary DNA was derived from fresh placenta; it was linked to the LTR of pRSVCat (2), also given by B. Howard.

RESULTS

A variety of genes has been cloned by others into selection vectors, but usually as a form of co-transformation because they are independently regulated. It is also possible to examine their promoter function in transient expression systems (2), but care has not always been given to the splice site and intervening nucleotides between the insert and reporter genes. Despite the caveats of semi-random integration, it was decided to stringently link gene activation and selection by the creation of an in-frame fusion of the 5'flank and first exon of human fetal globin and the coding region of a neomycin resistance gene (Fig. 1). In this construct, twelve amino acids of Neo are replaced by 14 similar amino acids of fetal globin, followed by an open reading frame as shown by DNA sequencing (Fig. 2). As the Neo gene can tolerate a deletion of similar size without loss of resistance (14), the hybrid protein should remain active. To remove uncertainty about active viral sequences, the SV40 promoter/enhancer region was deleted. Thus, one is left with a selectable reporter gene which must be driven by the fetal globin promoter for the cell to survive in the presence of G-418.

In this form, the vector $p^{G}\gamma$-Neo can simulate the activity of an endogenous gene when transfected into cells, and should confer an advantage where the parent gene is expressed but not ordinarily selectable.

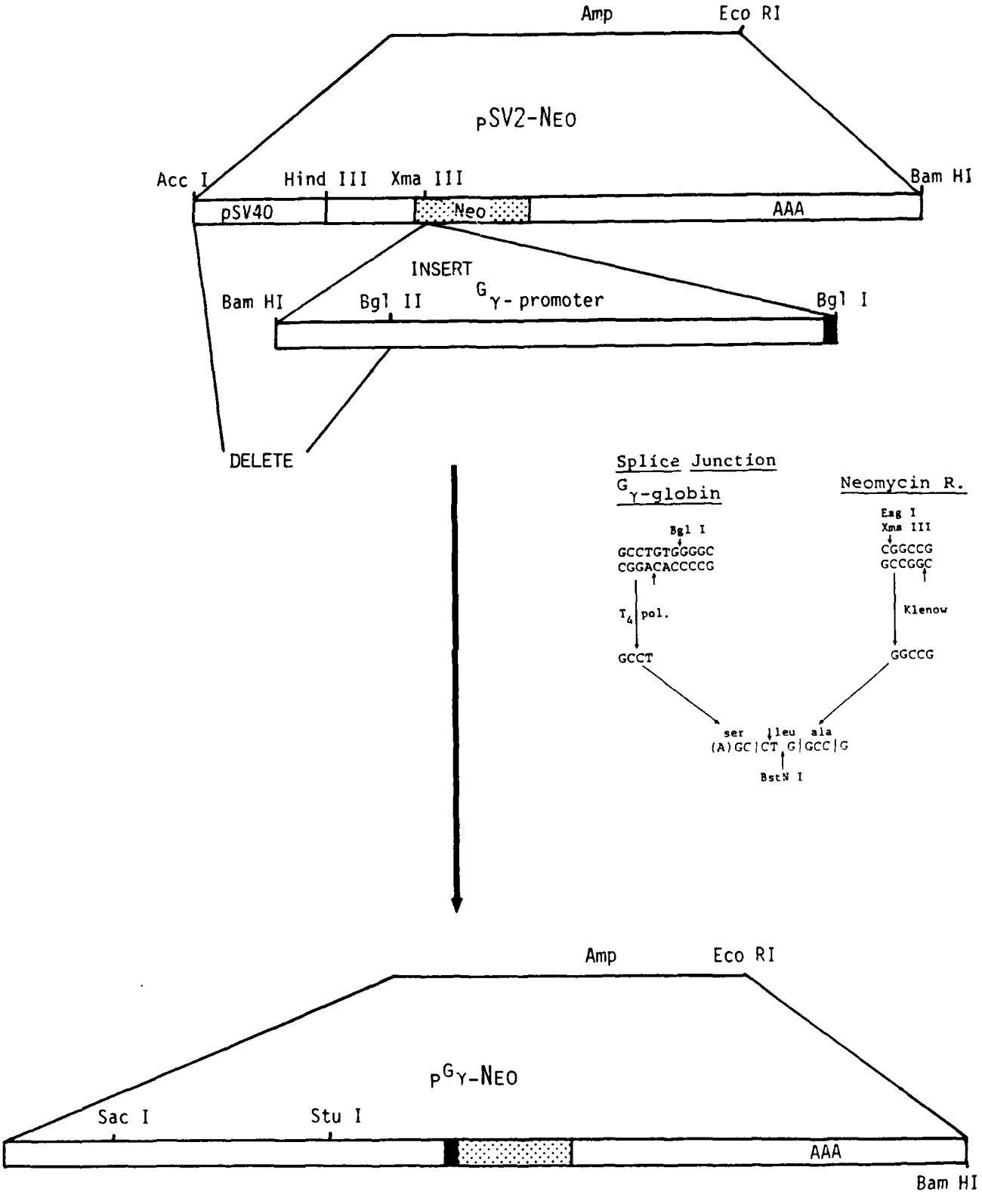

Figure 1. The insertion of a fetal hemoglobin promoter into the plasmid pSV2Neo (see Plasmid Construction in Methods). The globin-Neo splice junction is shown to the right of the arrow and the result of the fusion is at the bottom.

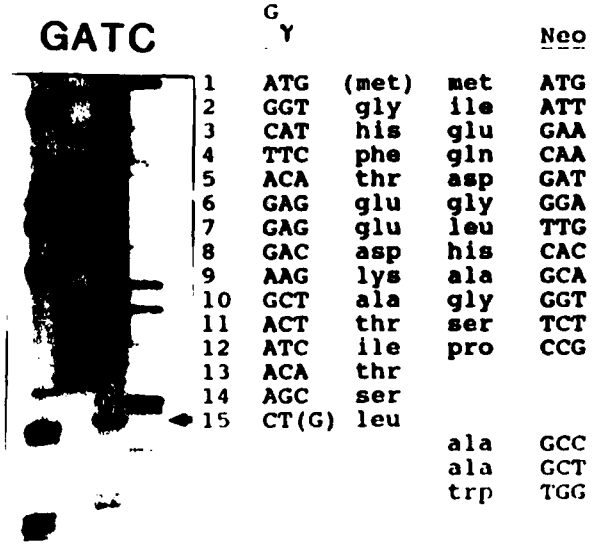

Figure 2. DNA sequence of the open reading frame between the first 14 codons of fetal hemoglobin and the 13th codon of the neomycin resistance gene as shown by the arrow.

To find activation factors, however, one must use host cells in which the promoter and reporter gene are quiescent, but which could be induced with complementary genes. This intermediate step can be accomplished by the introduction of a separate selectable marker, e.g. from pSV2Gpt (13) (Fig. 3). The vector contains the gene for the bacterial enzyme xanthine-guanine phosphoribosyltransferase, which can rescue cells metabolically blocked by aminopterin and mycophenolic acid.

Finally, the entire construction is transfected into permissive xenogeneic cells; in this case, a line derived from murine erythroleukemia. The cells are selected for growth in aminopterin and mycophenolic acid and individual clones are isolated. Their lack of growth in the neomycin analogue G-418 is confirmed and their DNA is analyzed to check for the presence of an unrearranged promoter-reporter gene. These clones were then propagated in culture.

The complementary DNA was prepared from a human placental library and linked to a retroviral long terminal repeat sequence as a promoter, and transfected by the calcium phosphate (15) or electroporation (16) techniques. The cells then were selected for the expression of resistance to G-418 (Fig.4). Clones were isolated, amplified in culture, and the DNA extracted for re-transfection of their parental cells in order to dilute out extraneous human sequences. Finally, clones were picked, grown, and their DNA was used to make another library. Digestion was carried out with restriction enzymes which would not disrupt the ß-lactamase gene, so that upon religation and transformation of bacteria, clones could be selected easily on ampicillin plates. In addition, the clones would be probed for the presence of human Alu-type repetitive sequences (17-19).

To establish the functional relevance of the putative activator clones, the sequences were used to probe mRNA from cells either making or not making $^{G}\gamma$-globin on Northern blots. Once confirmed, the sequences also could be used to pull out full-length clones from other DNA libraries if necessary. The ultimate test would be to transfect the activator genes into cells which do not contain the specific RNA and convert them into ones which express it.

DISCUSSION

The first part of this approach was mentioned in an obscure article in 1984 (20), where the selectable marker was Neo, driven by a thymidine kinase promoter, and the reporter gene was Gpt, regulated by a rat insulin promoter. That paper differed from the present method because the splice site occurred in the 5' untranslated region of Gpt, and that the entire construct was placed into a retrovirus. The latter step may not be necessary,

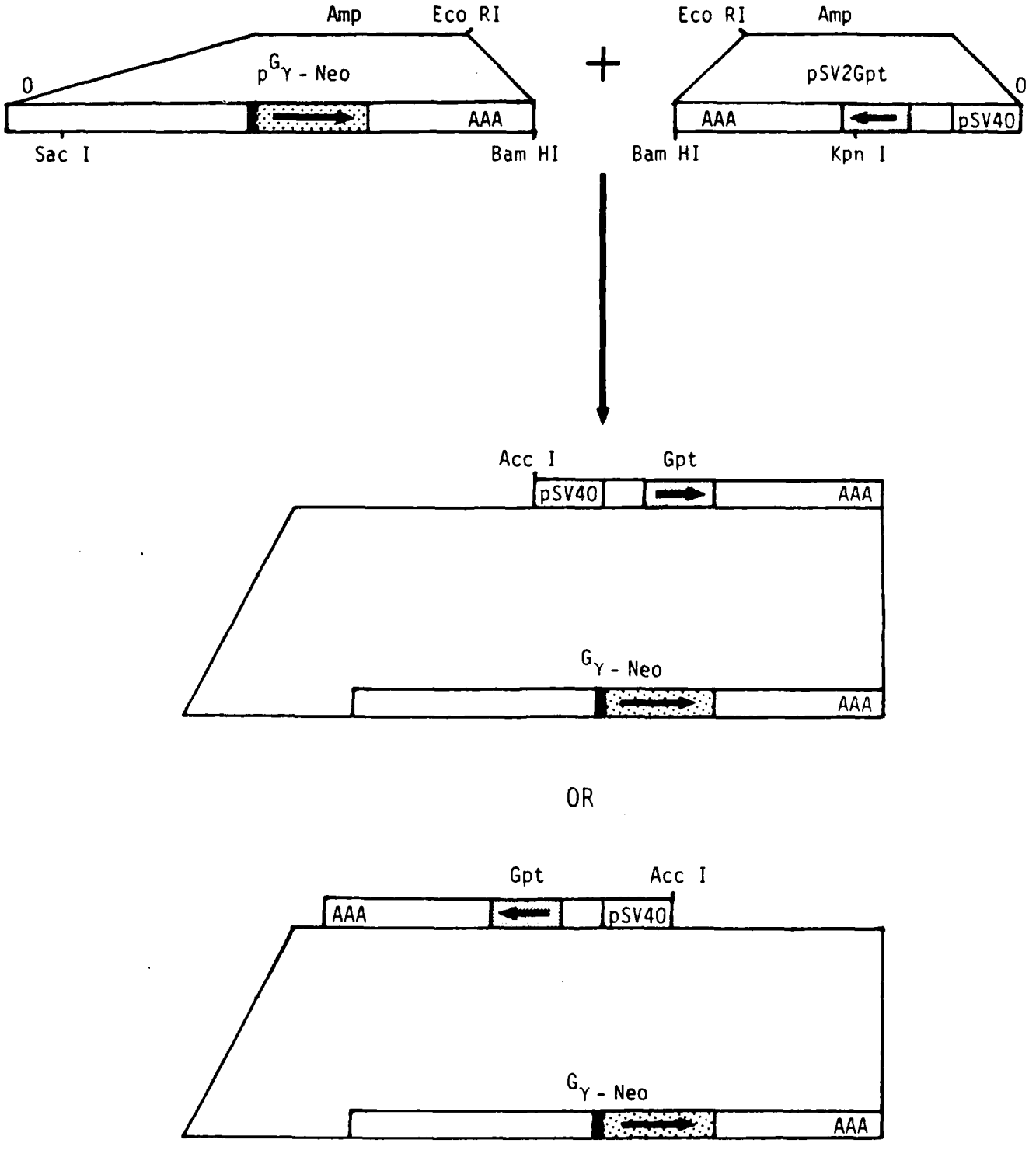

Figure 3. Introduction of a second selectable marker (Gpt). The construct in Fig. 1 and pSV2Gpt are spliced together, and as a result, their genes can be tandem or opposite in orientation.

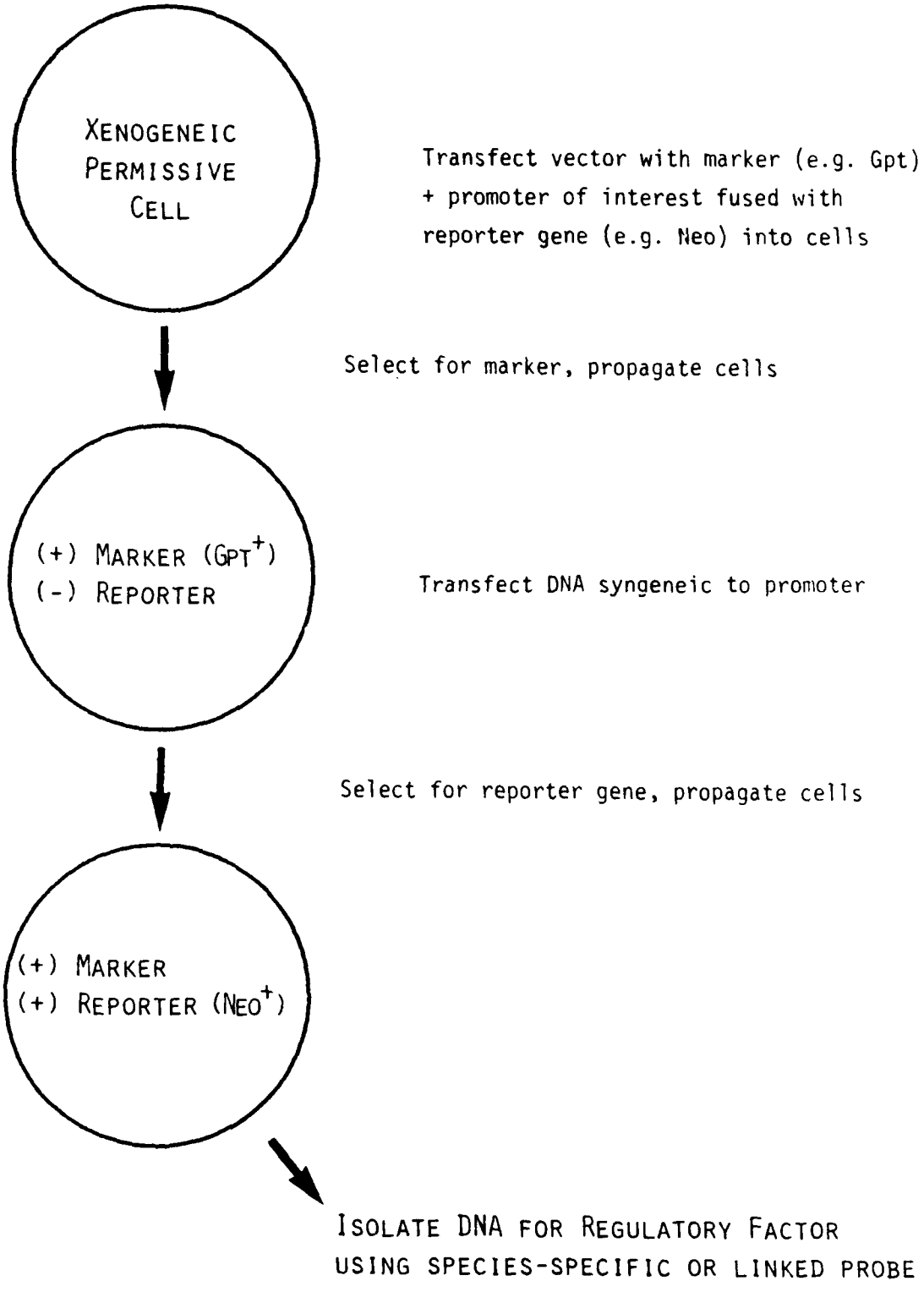

Figure 4. Overall strategy for the isolation of gene activation factors.

since only a moderate number of transformants are required, and indeed could produce problems because the LTR's could act as promoters to make anti-sense Gpt, which would impair the detection of Gpt(+) clones.

For the second phase of this experiment, the same authors developed another retroviral system to maximize the frequency of transfection (21). This vector had tri-initiation codons for expression of a cloned cDNA library and could be propagated as a bacteriophage in addition to a retrovirus. It also contained the selectable marker Neo, even though the host cells which it was designed to complement in the previous paper should already have contained this marker and it would make it more difficult to prove transformation than if a third marker were used. Although there is an advantage to the use of retroviruses for their high efficiency of transfection, their low copy numbers per cell will allow detection of only those activators where a single gene is sufficient to cause expression of the reporter. However, bulk transfection methods, (e.g. calcium phosphate), may allow several different regulators to interact. Then extraneous sequences can be diluted by subsequent rounds of transfection, as has been done to isolate oncogenes (17). In addition, as the retroviral sequences are integrated to the nucleus anyway, it is not certain what the advantage of cDNA is over the use of genomic DNA, where the introns could be excised naturally. Novel as this system was, it failed to work when applied to the insulin promoter (22).

The approach presented in this paper is similar, except for the differences noted above. Because retroviruses are not used here, the efficiency of transformation may be lower but compensated by bulk transfection methods and the use of xenogeneic permissive cells. The latter define a host in which all transcription factors, except for a single activator, are present. The source of these cells from a heterologous species facilitates the isolation and purification of unique human genes. Probably these strategies will not work in all systems because of repressors (and this issue will be addressed in a subsequent report). With luck, however, this approach will ultimately succeed. On that occasion, a major step will have been taken in Biology.

ACKNOWLEDGEMENTS

Naval Medical Research and Development Command, Work Unit No. 62233N.MM33C30.007.1003. The opinions and assertions contained herein are the private ones of the writer and are not to be construed as official or reflecting the views of the Navy Department or the Naval Service at large. The author wishes to thank Mrs. Lula Nicholas for assistance in preparing the manuscript.

REFERENCES

1. Maniatis T, Goodbourn T, Fischer JA: Regulation of inducible and tissue-specific gene expression. Science, 236:1237-1245, 1987.

2. Gorman CM, Merlino GT, Willingham MC, Pastan I, Howard BH: The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eukaryotic cells by DNA-mediated transfection. Proc Nat Acad Sci USA, 79:6777-6781, 1982.

3. Myers RM, Lerman LS, Maniatis T: A general method for saturation mutagenesis of cloned DNA fragments. Science, 229:242-247, 1985 (cf. ibid., 232:613-618, 1986).

4. Dignam JD, Lebovitz RM, Roeder RG: Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucl Acids Res, 11:1475-1489, 1983 (cf. Cell, 32:669-680, 1983).

5. Garner MM, Revzin A: A gel electrophoresis method for quantifying the finding of proteins to specific DNA regions: application to components of the E. coli lactose operon regulatory system. Nucl Acids Res, 9:3047-3060, 1981 (cf. Nucl Acids Res, 9:6505-6525, 1981).

6. Bowen B, Steinberg T, Laemmli UK, Weintraub H: The detection of DNA-binding proteins by protein blotting. Nucl Acids Res, 8:1-20, 1980. (cf. Gene Anal Tech, 1:3-8, 1984; Proc Nat Acad Sci USA, 82:6741-6744, 1985; Mol Cell Biol, 5:2967-2974, 1985).

7. Weintraub H, Groudine M: Chromosomal subunits in active genes have an altered conformation. Science, 193:848-856, 1976.

8. Galas D, Schmitz A: DNase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucl Acids Res, 5:3157-3170, 1978.

9. Jones KA, Kadonaga JT, Rosenfeld PJ, Kelly TJ, Tjian R: A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell, 48:79-89, 1987 (cf. Mol Cell Biol, 6:3117-3127, 1986; ibid., pp. 4723-4733).

10. Singh H, LeBowitz JH, Baldwin AS, Sharp PA: Molecular cloning of an enhancer binding protein: Isolation by screening of an expression library with a recognition site DNA. Cell, 52:415-423, 1988.

11. Maniatis T, Fritsch EF, Sambrook J: Molecular cloning. A laboratory manual, Cold Spring Harbor Laboratory, 1982.

12. Southern PJ, Berg P: Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J Mol Appl Gen, 1:327-341, 1982.

13. Mulligan RC, Berg P: Expression of a bacterial gene in mammalian cells. Science, 209:1422-1427, 1980.

14. Davies J.: personal communication.

15. Graham FL, van der Eb AJ: A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology, 52:456-467, 1973.

16. Wong TK, Neumann E: Electric field mediated gene transfer. Biochem Biophys Res Comm, 107:584-587, 1982.

17. Shih C, Weinberg RA: Isolation of a transforming sequence from a human bladder carcinoma cell line. Cell, 29:161-169, 1982.

18. Gusella JF, Keys C, Varsanyi-Breiner A, Kao F, Jones C, Puck TT, Housman D: Isolation and localization of DNA segments from specific human chromosomes. Proc Nat Acad Sci USA, 77:2829-2833, 1980 (cf. Cell, 29:161-169, 1982).

19. Broach JR, Strathern JN, Hicks JB: Transformation in yeast: Development of a hybrid cloning vector and isolation of the CAN1 gene. Gene, 8:121-133, 1979 (cf. Nature, 285:207-210, 1980).

20. Episkopou V, Murphy AJM, Efstratiadis A: Cell-specified expression of a selectable hybrid gene. Proc Nat Acad Sci USA, 81:4657-4661, 1984.

21. Murphy AJM, Efstratiadis A: Cloning vectors for expression of cDNA libraries in mammalian cells. Proc Nat Acad Sci, 84:8277-8281, 1987.

22. Efstratiadis A: personal communication.